

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 電気通信学研究科 電子工学専攻 博士前期課程		
氏 名	細川 嵩	学籍番号	0732079
論 文 題 目	局所解脱出を考慮した免疫型強化学習器に関する研究		
<p>要 旨</p> <p>生産工程などあらかじめ作業内容や環境が固定された状況で用いられる産業用ロボットに対し、最近では日常環境で用いられる家の中の掃除を行う家庭用ロボットや、介護用ロボット、人間の代わりに警備を行うロボットなどが数多く登場している。産業用ロボットなどでは目標や動作環境が固定されているので、最適制御問題などを解くことにより最適な行動を設定することができる。しかし、家庭用ロボットは運用先によって目標とする状態や目標到達に必要な行動セット政策が異なるため、それぞれの運用先に合わせた適切な政策を設定しなければならない。しかし、人間が個々のロボットに対してそれぞれ適切な政策を設定するのは大きな負荷となる。このため家庭用ロボットでは学習機構を搭載し、運用先に適したタスクを自動的に獲得させることが重要となる。</p> <p>ロボットに学習を行わせる手法として強化学習がある。強化学習は人間からの詳細な行動についての教示を必要としない学習法であり、人間は目標状態を設定するのみでよい。強化学習では事前に設定された行動テーブルから実行することのできる行動を試行錯誤的に選択し、環境から得られた報酬値を元に学習を行うため事前に情報のない環境でも自律的に学習を行わせることができる。しかし、学習初期では一般的にランダム探索となり、同じ状態を何度も探索してしまう恐れがあるため目標までの到達効率が良いとは言えない。また、多くの強化学習手法では環境変化に十分に対応できているとは言えず、環境変化によりロボットが局所解に陥ってしまい目標に到達できなくなってしまう場合がある。</p> <p>この局所解状態の脱出には生物の免疫の病原体に対して中和を行う抗体の濃度変化を利用した手法などがある。しかしこれらの局所解検出器は学習機構を搭載していないため目標方向などの環境情報を必要とし、ロボットが自律的に環境に適応した行動を獲得することができない。</p> <p>本研究では、強化学習器と局所解検出器を統合した学習器の構築を行うことにより学習初期の目標到達効率の改善、環境変化への対応を行う。提案する学習器はそれぞれ獲得免疫系を参考とした手法を用い、手法の一部を再モデル化することにより矛盾無く統合することが期待される。提案手法を既存の強化学習器では局所解状態に陥ることが多い環境変化のある迷路探索問題に適用し、その有効性を示す。</p>			